

Multi-Modal retinal image analysis via deep learning for the diagnosis of intermediate dry age-related macular degeneration: a feasibility study

Vaghefi, Ehsan; Hill, Sophie ; Kersten, Hannah M.; Squirrell , David

Published in:
Journal of Ophthalmology

DOI:
[10.1155/2020/7493419](https://doi.org/10.1155/2020/7493419)

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):
Vaghefi, E, Hill, S, Kersten, HM & Squirrell , D 2020, 'Multi-Modal retinal image analysis via deep learning for the diagnosis of intermediate dry age-related macular degeneration: a feasibility study', *Journal of Ophthalmology*, vol. 2020 , 7493419. <https://doi.org/10.1155/2020/7493419>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

Research Article

Multimodal Retinal Image Analysis via Deep Learning for the Diagnosis of Intermediate Dry Age-Related Macular Degeneration: A Feasibility Study

Ehsan Vaghefi ^{1,2}, Sophie Hill,³ Hannah M. Kersten,¹ and David Squirrel³

¹School of Optometry and Vision Sciences, University of Auckland, Auckland, New Zealand

²Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand

³Department of Ophthalmology, University of Auckland, Auckland, New Zealand

Correspondence should be addressed to Ehsan Vaghefi; e.vaghefi@auckland.ac.nz

Received 3 September 2019; Revised 28 November 2019; Accepted 19 December 2019; Published 13 January 2020

Academic Editor: Sentaro Kusuvara

Copyright © 2020 Ehsan Vaghefi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background and Objective. To determine if using a multi-input deep learning approach in the image analysis of optical coherence tomography (OCT), OCT angiography (OCT-A), and colour fundus photographs increases the accuracy of a CNN to diagnose intermediate dry age-related macular degeneration (AMD). **Patients and Methods.** Seventy-five participants were recruited and divided into three cohorts: young healthy (YH), old healthy (OH), and patients with intermediate dry AMD. Colour fundus photography, OCT, and OCT-A scans were performed. The convolutional neural network (CNN) was trained on multiple image modalities at the same time. **Results.** The CNN trained using OCT alone showed a diagnostic accuracy of 94%, whilst the OCT-A trained CNN resulted in an accuracy of 91%. When multiple modalities were combined, the CNN accuracy increased to 96% in the AMD cohort. **Conclusions.** Here we demonstrate that superior diagnostic accuracy can be achieved when deep learning is combined with multimodal image analysis.

1. Introduction

In the developed world, age-related macular degeneration (AMD) is the leading cause of irreversible vision loss in the population over 60 years old [1, 2]. Broadly AMD can be categorised in neovascular and non-neovascular AMD, with the later being far more prevalent. Characteristic features of non-neovascular AMD include macular drusen, RPE abnormalities, and, in the late stage, geographic atrophy [3]. Pigment abnormalities detected by colour fundus photography (CFP) are now well recognised to be one of the major risk factors for the development of late stage AMD [3–6].

Over time, optical coherence tomography (OCT) and fundus autofluorescence (FAF) have also joined the battery of imaging techniques that are now considered essential for the monitoring of non-neovascular AMD. This list has more recently been joined by optical coherence tomography angiography (OCT-A) and multicolour confocal scanning laser

ophthalmoscopy (SLO) [7]. With the introduction of OCT and OCT-A guided AMD treatment regimens, there has been an exponential increase in the number of images obtained and stored in large electronic databases [8]. Furthermore, the multiple imaging techniques that are now routinely used in clinical practice mean that the cumulative time spent by retinal specialists interpreting these images is also increasing [8]. Utilising software or computation power to automate image analysis may have the potential to improve the efficiency and accuracy of this process in the clinic [9]. Computer-assisted image assessment is also less prone to human factors such as bias, fatigue, and mindset [10]. Computer-assisted diagnosis is not a new concept: radiology adopted this approach when the increasing demand for imaging studies outstripped the ability of radiologists to interpret and report on the studies [11].

Within the field of ophthalmology, automated image analysis has been applied to the detection of diabetic

retinopathy, mapping visual field defects in glaucoma, and grading cataract [12–14]. Concentrating on macular disease, both semiautomated and automated techniques have already been used and validated in the detection of drusen, reticular drusen, and geographic atrophy [15–19]. The majority of these studies have developed networks using only a single imaging modality, namely, OCT [20–24], and very few have combined image data from more than one modality, e.g., CFP and OCT images [25], or infrared, green FAF, and SLO [26].

The purpose of this study is to determine whether a multimodal deep learning approach; training the CNN on OCT, OCT-A, and CFP, will diagnose intermediate dry AMD more accurately, when compared to conventional CNN trained on the single modalities of CFP, OCT, and OCT-A.

2. Methods

Seventy-five participants were recruited through Auckland Optometry Clinic and Milford Eye Clinic, Auckland, New Zealand. All participants provided written informed consent prior to imaging. Ethics approval (#018241) from the University of Auckland Human Participants Ethics Committee was obtained for this study. This research adhered to the tenets of the Declaration of Helsinki. Participants were divided into three groups: young healthy (YH), old healthy, (OH) and high-risk intermediate AMD (AMD). A total of 20 participants were recruited into the YH group, 21 into the OH group, and 34 into the AMD group. A comprehensive eye exam was conducted on each participant prior to the OCT and OCT-A scans including dilated fundal examination and high contrast best-corrected visual acuity (BCVA) to determine the ocular health status of the fundus. Patients with any posterior eye disease that could potentially affect the choroidal or retinal vasculature including but not limited to glaucoma, polypoidal choroidal vasculopathy, DR, hypertensive retinopathy, and high myopia (≥ 6 D) were excluded from the study. Patients with any history of neurological disorders were also excluded. None of our recruited participants fit the exclusion criteria. The YH group consisted entirely of individuals between the ages of 20 and 26 and a best corrected visual acuity of $\geq 6/9$ in the eye under test. The OH group consisted of individuals over the age of 55 years who had a best corrected visual acuity of $\geq 6/9$ in the eye under test and a normal ocular examination. The “AMD cohort” consisted entirely of patients with high-risk intermediate dry AMD. This was diagnosed if the individual had at least two of the following risk factors: reticular pseudodrusen, established neovascular AMD in the fellow eye, and confluent soft drusen with accompanying changes within the retinal pigment epithelium. In order to ensure that all patients in the “AMD cohort” could maintain fixation during OCT-A imaging, only those patients with a BCVA of 6/15 or better were enrolled. The mean age of the participants in the YH, OH, and AMD groups were 23 ± 3 , 65 ± 10 , and 75 ± 8 years, respectively. Only one eye of each patient was used for the analysis, and if the patient had both eyes scanned, the OCT-A scan that had the better quality

(assessed subjectively by the clinical grader) of the two was used. Mean BCVA for the YH, OH, and AMD groups were 6/6, 6/9, and 6/12, respectively. The ocular health of all participants was assessed at Auckland Optometry Clinic, by a registered optometrist, prior to enrolment in the study. The macular status of patients enrolled into the AMD group was assessed separately by an experienced retinal specialist (DS).

2.1. SS-OCT-A Device and Scanning Protocol. Participants were dilated with 1.0% tropicamide if the pupils were deemed too small for adequate OCT scans. The swept source (SS) OCT-A device (Topcon DRI OCT Triton, Topcon Corporation, Tokyo, Japan) was used to obtain the following images: $3 \times 3 \text{ mm}^2$ macular *en-face* OCT and $3 \times 3 \text{ mm}^2$ macular *en-face* OCT-A.

Raw OCT, OCT-A, and CFP image data were exported using Topcon IMAGENET 6.0 software. No image processing was performed prior to image analysis. The retinal layers were identified using the IMAGENET 6.0 automated layer detection tool (Figure 1). *En-face* OCT and OCT-A images from layers 6 to 9 of the scan were selected (ONL-IS, OS, RPE, and choroid), plus a single fundus photo. The dataset was divided into 0.6/0.2/0.2 for training, validation, and test sets. This division was based on participants (not images) in order to avoid data leakage between the training, validation, and test sets. As there were multiple OCT and OCT-A images per patient, appropriate measures were taken to ensure that there was no patient overlap between the training, validation, or test sets (Figure 2).

2.2. Convolutional Neural Network Design. The original design of the convolutional neural network (CNN) used here was based on the INCEPTION-RESNET-V2 design, since it appeared to have the advantage of faster convergence speed. This design was further modified to enable the network to be trained on multiple image modalities at the same time (Figure 3). Each imaging stream was then set up to initiate with a resizing layer, which was then followed by three repetitions of a 2D convolutional layer, batch normalization, and RELU activation layer. Separate modalities were then concatenated using a global pooling layer. The main body of the CNN followed the classic INCEPTION-RESNET-V2 design, where blocks of inception cells (A, B, and C) were used in series, where each block included cells with varying kernel sizes and channels, as well as feed forwards bypasses. The Python code for creating the CNN is released and freely available here: <https://medium.com/@mannasiladittya/building-inception-resnet-v2-in-keras-from-scratch-a3546c4d93f0>.

Experiments were run on an Intel Xeon Gold 6128 CPU @ 3.40 GHz with 16 GB of RAM memory and a NVIDIA GeForce TiTan V VOLTA 12 GB, for 100 epochs. The training-stop criterion was as CNN validation loss (measured as negative log-likelihood and residual sum of squares) had reached a stable minimum over the last 3 EPOCHs of training. Hence, any further training would have led to model “overfitting,” in which the neural network

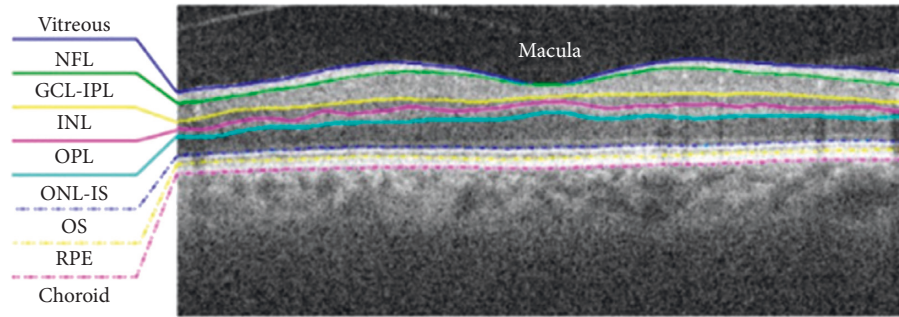


FIGURE 1: The retinal layers automatically identified by IMAGENet 6.0 software. NFL: nerve fiber layer; GCL: ganglion cell layer; IPL: inner plexiform layer; INL: inner nuclear layer; OPL: outer plexiform layer; ONL: outer nuclear layer; IS: inner segment; OS: outer segment; RPE: retinal pigment epithelium.

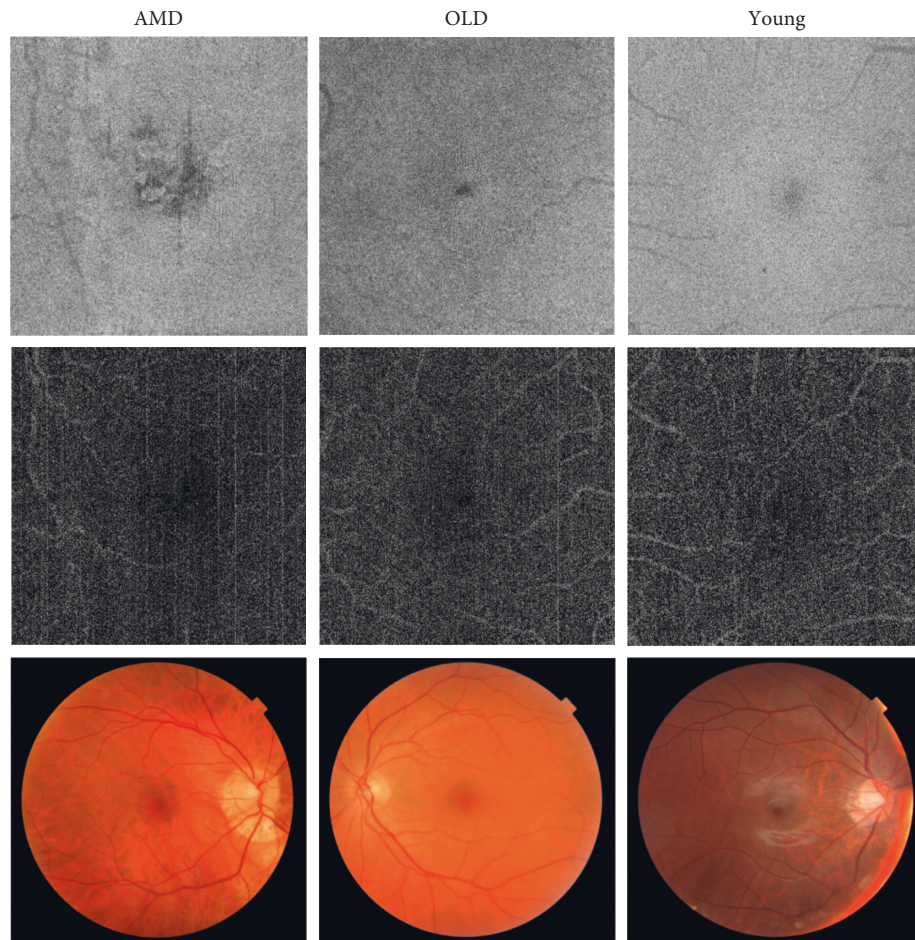


FIGURE 2: Examples of *en-face* OCT (top row), *en-face* OCT-A (middle row), and colour fundus photographs (bottom row) from each of the three cohorts (AMD, OH, and YH). The images within each column are from the same patient, from the labeled cohort at the top.

“memorizes” the training examples. It is worth noting that this approach is scalable to include more modalities (>3).

Several network training episodes were undertaken. Firstly, only a single image modality was used: OCT, OCT-A, or CFP. Further training episodes combined both OCT and OCT-A image data. The last training episode utilised multimodal image analysis from OCT, OCT-A, and CFP.

To better understand the nature of “learning” from our CNN in each modality, attention maps were produced from

each modality. This method is explained in detail elsewhere [27–29]. Briefly, the output of the last convolutional layer prior to the global concatenation was extracted, resized, and smoothed for visualization.

3. Results

Table 1 shows that the sensitivity and specificity of the CNNs trained using a single modality are high. However, when

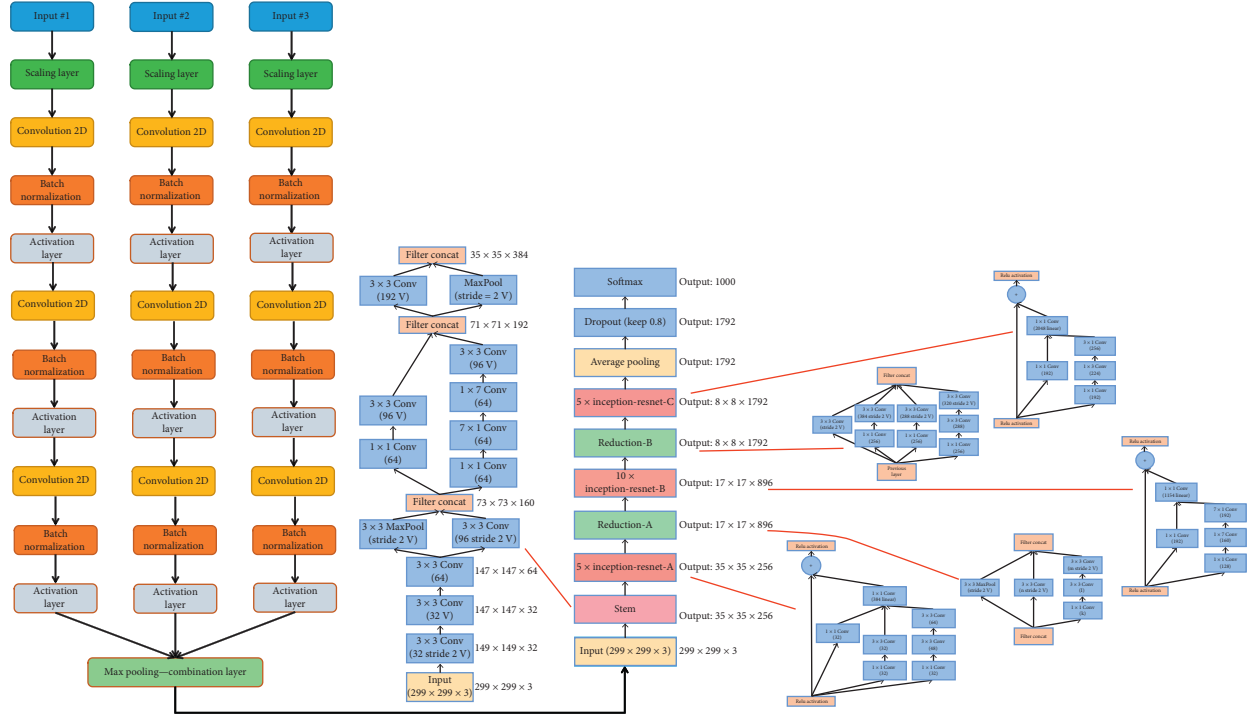


FIGURE 3: Modification to the top layers of INCEPTION-RESNET-V2 architecture.

TABLE 1: The accuracy, sensitivity, and specificity of each CNN design (single modality vs. dual modality vs. multimodality) in classification of three cohorts of YH, OH, and AMD.

	Patient cohort		
	YH	OH	AMD
Single modality CNN	Overall accuracy 94.4%		
OCT sensitivity (%)	99.6	98.9	77.8
OCT specificity (%)	98.8	86.7	100
Single modality CNN	Overall accuracy 91.9%		
OCT-A sensitivity (%)	95.5	83.2	97.6
OCT-A specificity (%)	99.6	96.2	76.4
Single modality CNN	Overall accuracy 93.8%		
CFP sensitivity (%)	81.0	84.6	100
CFP specificity (%)	94.4	78.6	86.7
Dual modality CNN (OCT + OCT-A)	Overall accuracy 96.7%		
Sensitivity (%)	100	95.7	92.1
Specificity (%)	97.6	94.6	98.7
Dual modality CNN (OCT + CFP)	Overall accuracy 92.9%		
Sensitivity (%)	98.1	96.3	100
Specificity (%)	100	97.0	96.0
Multimodality CNN (OCT + OCT-A + CFP)	Overall accuracy 99.8%		
Sensitivity (%)	100	99.3	100
Specificity (%)	100	100	99.2

more than one image modality is used during the training of the CNN, the sensitivity and specificity increase with each

additional image modality added. The CNN with the highest accuracy, sensitivity, and specificity was the multimodal image-trained network.

If the CNN results are considered for each imaging modality, it would suggest that each modality is better suited for classifying certain categories. The CFP single modality CNN was the most sensitive to AMD closely followed by OCT-A. In contrast, the single modality OCT was more sensitive to ageing, identifying the young and old cohorts more accurately. Combining the imaging modalities into a single “multimodal” CNN resulted in an accuracy of 99.8%, being able to identify both ageing and disease with high sensitivity and specificity (Table 1).

To investigate the apparent different sensitivities for each modality, “attention maps” of each modality were generated (Figure 4) from the same images in Figure 2. These maps demonstrate the image features that were “noticed” by multiple CNNs for each modality or participant group.

The attention maps of the OCT images in each cohort pay highest attention to the background homogeneity. In the OCT-A images, the areas of highest attention are what appear to be projection artefacts of the retinal vessels seen within the choriocapillaris OCTA slab, and the surrounding background OCTA signal of the adjacent tissues. In CFP, the area of highest attention is at the optic disc and the peripapillary region in young and AMD and additionally at the macula in the old cohort.

4. Discussion

In the current study, we wanted to compare the accuracy of different CNN designs, trained on the same dataset, to

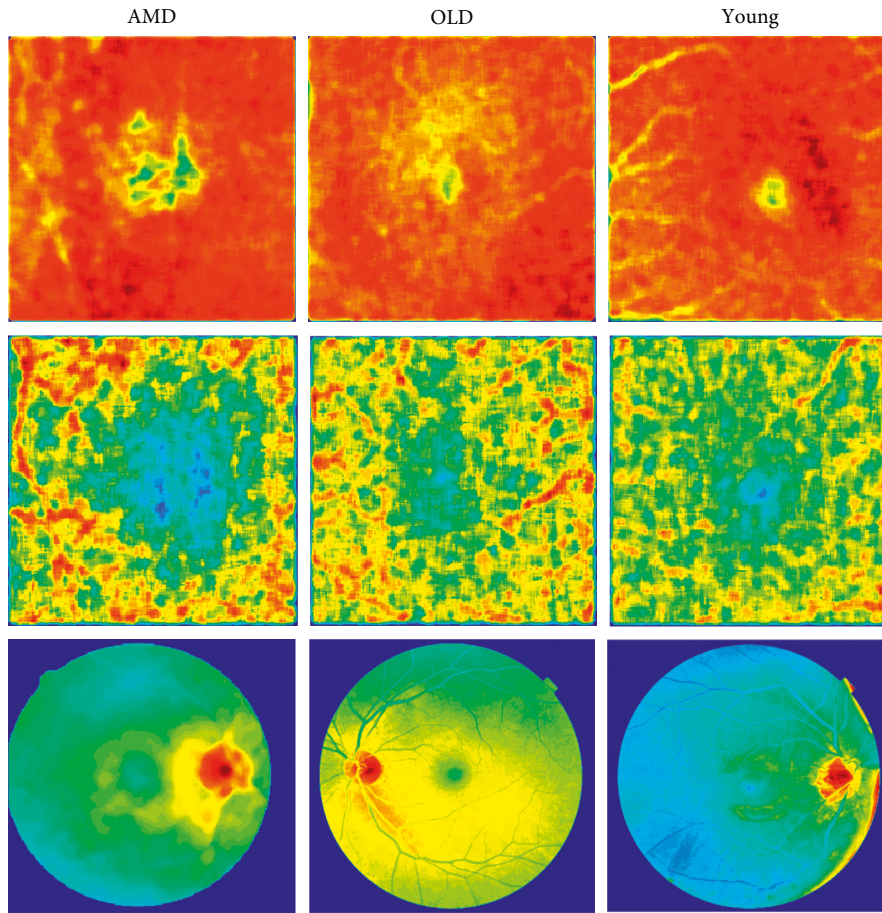


FIGURE 4: Same images as Figure 2 were used to generate these attention maps. Normalized images [0-1] are pseudocoloured based on the level of “attention” the CNN gave to the region in the image modality (low represented by blue, to high represented by red).

investigate whether combining imaging modalities improved the ability of the CNN to accurately identify the 3 distinct clinical groups under test. A secondary aim was to evaluate the attention maps of each design to investigate which components of the image(s), the CNN “paid attention” to.

To date, a number of single image modality CNNs designed to identify AMD have been published, with differing rates of success [30–36]. Furthermore, the majority of these studies have been based on single transverse OCT image (cross section) taken through the fovea [23, 37–40]. Although such an approach, training a CNN on a single OCT image, can achieve impressive results, this approach is flawed as a single transverse image will only sample a very limited part of the macula and relies on the pathology being present within the scan analysed.

The use of *en-face* OCT images may overcome the segmentation error and sampling bias associated with the use of transverse OCT scans. This technique has previously been utilised by Zhao et al. [33], to automatically detect the presence of drusen, including pseudodrusen. CFPs have also been used as a single imaging modality in the identification of AMD using an appropriately trained CNN [30, 32, 34–36, 41]. Good levels of accuracy have been obtained [30, 34], but again often only after image resizing and significant image preprocessing [36].

To the best of our knowledge, this is the first study to develop a CNN that utilises both *en-face* OCT and OCT-A and then combines them with CFP to develop a truly “multimodal” CNN, one that truly samples the entirety of the macula under test. To avoid sampling bias, the *en-face* OCT and OCT-A data slabs of the entire macula were used. We found that the multimodal CNN was superior to the single modality CNNs; moreover, incorporating additional modalities led to an incremental improved accuracy of the CNN.

To the best of our knowledge, only two other groups have utilised a similar multimodal approach [25, 42]. Yoo et al. [25] used paired fundus and OCT images, employing a VGG19 model pretrained on ImageNet to extract visual features from both images. More impressive is the approach taken by Wang et al. [42], who used a training method they termed “loose pairing” whereby pairings were constructed based on labels instead of eyes. In this method, a fundus image is allowed to be paired with an OCT image if their labels are identical, an approach which yielded an overall accuracy of 97%. The progressive improvement in the performance of the incrementally complex CNNs that Wang et al. [42] and ourselves describe strongly suggest that, like clinicians, the accuracy of a CNN to detect pathology and “normal” ageing is enhanced if complimentary imaging modalities are employed.

The outcome of our single modality CNNs (Table 1) suggest that OCT and OCT-A modalities are inherently sensitive to different aspects of the retinal health. It appears that OCT is more sensitive to ageing of the retina, while CFP and OCT-A are better suited to detect the pathological changes attributable to AMD. The black box nature of the neural networks makes any interpretation of the results problematic [43]. We therefore produced attention maps to aid our understanding of the CNN activity. The *en-face* OCT images in each cohort revealed that the highest attention was directed to the background homogeneity with a lesser emphasis on the fovea. OCT-A in contrast appears to be more sensitive to disease. Review of the attention maps reveals that within the OCT-A images, the regions of higher attention are the retinal vessels and the tissues immediately adjacent to them. Again the fovea appears to contribute very little to the OCT-A CNN classifier.

In conclusion, although trained on a small number of images, this study demonstrates that, compared with CNNs trained on a single image modality, superior diagnostic accuracy can be achieved when deep learning is combined with multimodal image analysis. We should also emphasise that this is only a “proof of concept” study, and larger studies to test its validity are needed. It should also be noted that the attention maps produced here are from our small cohort and may not prove to be generalisable after further studies. This approach is more similar to the multimodal image interpretation used by retinal specialists within the clinical environment, and our results suggest that this approach warrants further investigation. The limitations of this study include the small sample size of images, with the CNN trained using images from a single academic center. These clearly limit the generalisability of the CNNs we have trained, but arguably do not detract from the conclusion that a multimodal CNN is superior to a single modality CNN. Further research would include the utilisation of images from other OCT manufacturers, and validation of this CNN on a dataset from another academic or clinical institution.

Data Availability

The ophthalmic imaging data used to support the findings of this study were obtained under appropriate approval by the University of Auckland Ethics Committee and so cannot be made freely available. Requests for access to these data should be made to the corresponding author Dr Ehsan Vaghefi e.vaghefi@auckland.ac.nz, which will be then passed on to the University of Auckland Ethics Committee for further process.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Ministry of Business, Innovation and Education (MBIE) Endeavour Fast Start Fund (UOAX1805).

References

- [1] R. Klein, B. E. K. Klein, and K. L. P. Linton, “Prevalence of age-related maculopathy,” *Ophthalmology*, vol. 99, no. 6, pp. 933–943, 1992.
- [2] R. Klein, K. J. Cruickshanks, S. D. Nash et al., “The prevalence of age-related macular degeneration and associated risk factors,” *Archives of Ophthalmology*, vol. 128, no. 6, pp. 750–758, 2010.
- [3] F. L. Ferris, M. D. Davis, T. E. Clemons et al., “A simplified severity scale for age-related macular degeneration,” *Archives of Ophthalmology*, vol. 123, no. 11, pp. 1570–1574, 2005.
- [4] M. L. Klein, F. L. Ferris 3rd, J. Armstrong et al., “Retinal precursors and the development of geographic atrophy in age-related macular degeneration,” *Ophthalmology*, vol. 115, no. 6, pp. 1026–1031, 2008.
- [5] J. J. Wang, E. Rochtchina, A. J. Lee et al., “Ten-year incidence and progression of age-related maculopathy: the blue mountains eye study,” *Ophthalmology*, vol. 114, no. 1, pp. 92–98, 2007.
- [6] R. Klein, B. E. K. Klein, M. D. Knudtson, S. M. Meuer, M. Swift, and R. E. Gangnon, “Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver dam eye study,” *Ophthalmology*, vol. 114, no. 2, pp. 253–262, 2007.
- [7] S. T. Garrity, D. Sarraf, K. B. Freund, and S. R. Sadda, “Multimodal imaging of nonneovascular age-related macular degeneration,” *Investigative Ophthalmology & Visual Science*, vol. 59, no. 4, pp. AMD48–AMD64, 2018.
- [8] C. S. Lee, D. M. Baughman, and A. Y. Lee, “Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration,” *Ophthalmology Retina*, vol. 1, no. 4, pp. 322–327, 2017.
- [9] M. D. Abramoff, M. K. Garvin, and M. Sonka, “Retinal imaging and image analysis,” *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.
- [10] S. Schmitz-Valckenberg, A. P. Göbel, S. C. Saur et al., “Automated retinal image analysis for evaluation of focal hyperpigmentary changes in intermediate age-related macular degeneration,” *Translational Vision Science & Technology*, vol. 5, no. 2, p. 3, 2016.
- [11] F. Smielauskas, H. MacMahon, R. Salgia, and Y.-C. T. Shih, “Geographic variation in radiologist capacity and widespread implementation of lung cancer CT screening,” *Journal of Medical Screening*, vol. 21, no. 4, pp. 207–215, 2014.
- [12] M. D. Abramoff, Y. Lou, A. Erginay et al., “Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning,” *Investigative Ophthalmology & Visual Science*, vol. 57, no. 13, pp. 5200–5206, 2016.
- [13] R. Asaoka, H. Murata, A. Iwase, and M. Araie, “Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier,” *Ophthalmology*, vol. 123, no. 9, pp. 1974–1980, 2016.
- [14] X. Gao, S. Lin, and T. Y. Wong, “Automatic feature learning to grade nuclear cataracts based on deep learning,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2693–2701, 2015.
- [15] M. J. J. P. van Grinsven, Y. T. E. Lechanteur, J. P. H. van de Ven et al., “Automatic drusen quantification and risk assessment of age-related macular degeneration on color fundus images,” *Investigative Ophthalmology & Visual Science*, vol. 54, no. 4, pp. 3019–3027, 2013.

- [16] Q. Chen, T. Leng, L. Zheng et al., "Automated drusen segmentation and quantification in SD-OCT images," *Medical Image Analysis*, vol. 17, no. 8, pp. 1058–1072, 2013.
- [17] A. D. Mora, P. M. Vieira, A. Manivannan, and J. M. Fonseca, "Automated drusen detection in retinal images using analytical modelling algorithms," *BioMedical Engineering OnLine*, vol. 10, no. 1, p. 59, 2011.
- [18] V. Sivagnanavel, R. T. Smith, G. B. Lau, J. Chan, C. Donaldson, and N. V. Chong, "An interinstitutional comparative study and validation of computer aided drusen quantification," *British Journal of Ophthalmology*, vol. 89, no. 5, pp. 554–557, 2005.
- [19] S. Schmitz-Valckenberg, C. K. Brinkmann, F. Alten et al., "Semiautomated image processing method for identification and quantification of geographic atrophy in age-related macular degeneration," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 10, pp. 7640–7646, 2011.
- [20] R. Poplin, A. V. Varadarajan, K. Blumer et al., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, pp. 158–164, 2018.
- [21] J. Krause, V. Gulshan, E. Rahimy et al., "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy," *Ophthalmology*, vol. 125, no. 8, pp. 1264–1272, 2018.
- [22] J. De Fauw, J. R. Ledsam, B. Romera-Paredes et al., "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature Medicine*, vol. 24, no. 9, pp. 1342–1350, 2018.
- [23] D. S. Kermany, M. Goldbaum, W. Cai et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131.e9, 2018.
- [24] R. F. Mansour, "Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy," *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 41–57, 2018.
- [25] T. K. Yoo, J. Y. Choi, J. G. Seo, B. Ramasubramanian, S. Selvaperumal, and D. W. Kim, "The possibility of the combination of OCT and fundus images for improving the diagnostic accuracy of deep learning for age-related macular degeneration: a preliminary experiment," *Medical & Biological Engineering & Computing*, vol. 57, no. 3, pp. 677–687, 2019.
- [26] A. Ly, L. Nivison-Smith, N. Assaad, and M. Kalloniatis, "Multispectral pattern recognition reveals a diversity of clinical signs in intermediate age-related macular degeneration," *Investigative Ophthalmology & Visual Science*, vol. 59, no. 5, pp. 1790–1799, 2018.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, Curran Associates, Inc., New York, NY, USA, 2012.
- [28] K. Cho, A. Courville, and Y. Bengio, "Describing multimedia content using attention-based encoder-decoder networks," *IEEE Transactions on Multimedia*, vol. 17, no. 11, pp. 1875–1886, 2015.
- [29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, <https://arxiv.org/abs/1409.0473>.
- [30] F. Grassmann, J. Mengelkamp, C. Brandl et al., "A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography," *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, 2018.
- [31] M. Treder, J. L. Laueremann, and N. Eter, "Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning," *Graefes' Archive for Clinical and Experimental Ophthalmology*, vol. 256, no. 2, pp. 259–265, 2018.
- [32] P. M. Burlina, N. Joshi, K. D. Pacheco, D. E. Freund, J. Kong, and N. M. Bressler, "Use of deep learning for detailed severity characterization and estimation of 5-year risk among patients with age-related macular degeneration," *JAMA Ophthalmology*, vol. 136, no. 12, pp. 1359–1366, 2018.
- [33] R. Zhao, A. Camino, J. Wang et al., "Automated drusen detection in dry age-related macular degeneration by multiple-depth, en face optical coherence tomography," *Biomedical Optics Express*, vol. 8, no. 11, pp. 5049–5064, 2017.
- [34] J. H. Tan, S. V. Bhandary, S. Sivaprasad et al., "Age-related macular degeneration detection using deep convolutional neural network," *Future Generation Computer Systems*, vol. 87, pp. 127–135, 2018.
- [35] U. R. Acharya, Y. Hagiwara, J. E. W. Koh et al., "Automated screening tool for dry and wet age-related macular degeneration (ARMD) using pyramid of histogram of oriented gradients (PHOG) and nonlinear features," *Journal of Computational Science*, vol. 20, pp. 41–51, 2017.
- [36] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, "Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks," *JAMA Ophthalmology*, vol. 135, no. 11, pp. 1170–1176, 2017.
- [37] D. B. Russakoff, A. Lamin, J. D. Oakley, A. M. Dubis, and S. Sivaprasad, "Deep learning for prediction of AMD progression: a pilot study," *Investigative Ophthalmology & Visual Science*, vol. 60, no. 2, pp. 712–722, 2019.
- [38] P. Serrano-Aguilar, R. Abreu, L. Antón-Canalis et al., "Development and validation of a computer-aided diagnostic tool to screen for age-related macular degeneration by optical coherence tomography," *British Journal of Ophthalmology*, vol. 96, no. 4, pp. 503–507, 2012.
- [39] W. Sun, X. Liu, and Z. Yang, "Automated detection of age-related macular degeneration in OCT images using multiple instance learning," in *Proceedings of the 9th International Conference on Digital Image Processing (ICDIP 2017)*, Hong Kong, China, May 2017.
- [40] Y. Wang, Y. Zhang, Z. Yao, R. Zhao, and F. Zhou, "Machine learning based detection of age-related macular degeneration (AMD) and diabetic macular edema (DME) from optical coherence tomography (OCT) images," *Biomedical Optics Express*, vol. 7, no. 12, pp. 4928–4940, 2016.
- [41] Y. Zheng, B. Vanderbeek, E. Daniel et al., "An automated drusen detection system for classifying age-related macular degeneration with color fundus photographs," in *Proceedings of the 2013 IEEE 10th International Symposium on Biomedical Imaging*, pp. 1448–1451, IEEE, San Francisco, CA, USA, April 2013.
- [42] W. Wang, Z. Xu, W. Yu et al., "Two-stream CNN with loose pair training for multi-modal AMD categorization," 2019, <https://arxiv.org/abs/1907.12023>.
- [43] D. Castellevecchi, "Can we open the black box of AI?," *Nature*, vol. 538, no. 7623, pp. 20–23, 2016.

